AD NO. ______________________
DTC PROJECT NO. 8-CO-160-UXO-021
REPORT NO. ATC-9109

STANDARDIZED

UXO TECHNOLOGY DEMONSTRATION SITE

OPEN FIELD SCORING RECORD NO. 298

SITE LOCATION:
U.S. ARMY ABERDEEN PROVING GROUND

DEMONSTRATOR:
GEO-CENTERS, INC.
7 WELLS AVENUE
NEWTON, MA 02459

TECHNOLOGY TYPE/PLATFORM:
SIMULTANEOUS EM AND MAGNETOMETRY
(MULTISENSOR STOLS)/TOWED ARRAY

PREPARED BY:
U.S. ARMY ABERDEEN TEST CENTER
ABERDEEN PROVING GROUND, MD 21005-5059

OCTOBER 2005

Prepared for:
U.S. ARMY ENVIRONMENTAL CENTER
ABERDEEN PROVING GROUND, MD 21010-5401

U.S. ARMY DEVELOPMENTAL TEST COMMAND
ABERDEEN PROVING GROUND, MD 21005-5055

DISPOSITION INSTRUCTIONS

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
October 2005	Final	4 through 6 August 2004

4. TITLE AND SUBTITLE
STANDARDIZED UXO TECHNOLOGY DEMONSTRATION SITE OPEN FIELD SCORING RECORD NO. 298 (GEO-CENTERS, INC.)

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Overbay, Larry; Robitaille, George
The Standardized UXO Technology Demonstration Site Scoring Committee

5d. PROJECT NUMBER
8-CO-160-UXO-021

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Aberdeen Test Center
ATTN: CSTE-DTC-AT-SL-E
Aberdeen Proving Ground, MD 21005-5059

8. PERFORMING ORGANIZATION REPORT NUMBER
ATC-9109

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Environmental Center
ATTN: SFIM-AEC-ATT
Aberdeen Proving Ground, MD 21005-5401

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
Same as Item 8

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This scoring record documents the efforts of GEO-CENTERS, Inc., to detect and discriminate inert unexploded ordnance (UXO) utilizing the APG Standardized UXO Technology Demonstration Site Open Field. Scoring Records have been coordinated by Larry Overbay and the Standardized UXO Technology Demonstration Site Scoring Committee. Organizations on the committee include, the U.S. Army Corps of Engineers, the Environmental Security Technology Certification Program, the Strategic Environmental Research and Development Program, the Institute for Defense Analysis, the U.S. Army Environmental Center, and the U.S. Army Aberdeen Test Center.

15. SUBJECT TERMS
GEO-CENTERS Inc., UXO Standardized Technology Demonstration Site Program, Open Field, Simultaneous EM and Magnetometry (Multi-Sensor STOLS)/towed array

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(Include area code)*
Unclassified	Unclassified	Unclassified	UL		

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

ACKNOWLEDGEMENTS

Authors:

Larry Overbay Jr.
Matthew Boutin
Military Environmental Technology Demonstration Center (METDC)
U.S. Army Aberdeen Test Center (ATC)
U.S. Army Aberdeen Proving Ground (APG)

Rick Fling
Aberdeen Test and Support Services (ATSS)
Sverdrup Technology, Inc.
U.S. Army Aberdeen Proving Ground (APG)

Christina McClung
Aberdeen Data Services Team (ADST)
Tri-S, Inc.
U.S. Army Aberdeen Proving Ground (APG)

Contributor:

George Robitaille
U.S. Army Environmental Center (AEC)
U.S. Army Aberdeen Proving Ground (APG)

TABLE OF CONTENTS

SECTION 4. TECHNICAL PERFORMANCE RESULTS

SECTION 5. ON-SITE LABOR COSTS

SECTION 6. COMPARISON OF RESULTS TO BLIND GRID DEMONSTRATION

SECTION 7. APPENDIXES

SECTION 1. GENERAL INFORMATION

1.1 BACKGROUND

Technologies under development for the detection and discrimination of unexploded ordnance (UXO) require testing so that their performance can be characterized. To that end, Standardized Test Sites have been developed at Aberdeen Proving Ground (APG), Maryland and U.S. Army Yuma Proving Ground (YPG), Arizona. These test sites provide a diversity of geology, climate, terrain, and weather as well as diversity in ordnance and clutter. Testing at these sites is independently administered and analyzed by the government for the purposes of characterizing technologies, tracking performance with system development, comparing performance of different systems, and comparing performance in different environments.

The Standardized UXO Technology Demonstration Site Program is a multi-agency program spearheaded by the U.S. Army Environmental Center (AEC). The U.S. Army Aberdeen Test Center (ATC) and the U.S. Army Corps of Engineers Engineering Research and Development Center (ERDC) provide programmatic support. The program is being funded and supported by the Environmental Security Technology Certification Program (ESTCP), the Strategic Environmental Research and Development Program (SERDP) and the Army Environmental Quality Technology Program (EQT).

1.2 SCORING OBJECTIVES

The objective in the Standardized UXO Technology Demonstration Site Program is to evaluate the detection and discrimination capabilities of a given technology under various field and soil conditions. Inert munitions and clutter items are positioned in various orientations and depths in the ground.

The evaluation objectives are as follows:

a. To determine detection and discrimination effectiveness under realistic scenarios that vary targets, geology, clutter, topography, and vegetation.

b. To determine cost, time, and manpower requirements to operate the technology.

c. To determine demonstrator's ability to analyze survey data in a timely manner and provide prioritized "Target Lists" with associated confidence levels.

d. To provide independent site management to enable the collection of high quality, ground-truth, geo-referenced data for post-demonstration analysis.

1.2.1 Scoring Methodology

a. The scoring of the demonstrator's performance is conducted in two stages. These two stages are termed the RESPONSE STAGE and DISCRIMINATION STAGE. For both stages, the probability of detection (P_d) and the false alarms are reported as receiver-operating characteristic (ROC) curves. False alarms are divided into those anomalies that correspond to emplaced clutter items, measuring the probability of false positive (P_{fp}), and those that do not correspond to any known item, termed background alarms.

b. The RESPONSE STAGE scoring evaluates the ability of the system to detect emplaced targets without regard to ability to discriminate ordnance from other anomalies. For the blind grid RESPONSE STAGE, the demonstrator provides the scoring committee with a target response from each and every grid square along with a noise level below which target responses are deemed insufficient to warrant further investigation. This list is generated with minimal processing and, since a value is provided for every grid square, will include signals both above and below the system noise level.

c. The DISCRIMINATION STAGE evaluates the demonstrator's ability to correctly identify ordnance as such and to reject clutter. For the blind grid DISCRIMINATION STAGE, the demonstrator provides the scoring committee with the output of the algorithms applied in the discrimination-stage processing for each grid square. The values in this list are prioritized based on the demonstrator's determination that a grid square is likely to contain ordnance. Thus, higher output values are indicative of higher confidence that an ordnance item is present at the specified location. For digital signal processing, priority ranking is based on algorithm output. For other discrimination approaches, priority ranking is based on human (subjective) judgment. The demonstrator also specifies the threshold in the prioritized ranking that provides optimum performance, (i.e. that is expected to retain all detected ordnance and rejects the maximum amount of clutter).

d. The demonstrator is also scored on EFFICIENCY and REJECTION RATIO, which measures the effectiveness of the discrimination stage processing. The goal of discrimination is to retain the greatest number of ordnance detections from the anomaly list, while rejecting the maximum number of anomalies arising from non-ordnance items. EFFICIENCY measures the fraction of detected ordnance retained after discrimination, while the REJECTION RATIO measures the fraction of false alarms rejected. Both measures are defined relative to performance at the demonstrator-supplied level below which all responses are considered noise, i.e., the maximum ordnance detectable by the sensor and its accompanying false positive rate or background alarm rate.

e. Based on configuration of the ground truth at the standardized sites and the defined scoring methodology, there exists the possibility of having anomalies within overlapping halos and/or multiple anomalies within halos. In these cases, the following scoring logic is implemented:

(1) In situations where multiple anomalies exist within a single R_{halo}, the anomaly with the strongest response or highest ranking will be assigned to that particular ground truth item.

(2) For overlapping R_{halo} situations, ordnance has precedence over clutter. The anomaly with the strongest response or highest ranking that is closest to the center of a particular ground truth item gets assigned to that item. Remaining anomalies are retained until all matching is complete.

(3) Anomalies located within any R_{halo} that do not get associated with a particular ground truth item are thrown out and are not considered in the analysis.

f. All scoring factors are generated utilizing the Standardized UXO Probability and Plot Program, version 3.1.1.

1.2.2 Scoring Factors

Factors to be measured and evaluated as part of this demonstration include:

a. Response Stage ROC curves:

(1) Probability of Detection (P_d^{res}).

(2) Probability of False Positive (P_{fp}^{res}).

(3) Background Alarm Rate (BAR^{res}) or Probability of Background Alarm (P_{BA}^{res}).

b. Discrimination Stage ROC curves:

(1) Probability of Detection (P_d^{disc}).

(2) Probability of False Positive (P_{fp}^{disc}).

(3) Background Alarm Rate (BAR^{disc}) or Probability of Background Alarm (P_{BA}^{disc}).

c. Metrics:

(1) Efficiency (E).

(2) False Positive Rejection Rate (R_{fp}).

(3) Background Alarm Rejection Rate (R_{BA}).

d. Other:

(1) Probability of Detection by Size and Depth.

(2) Classification by type (i.e., 20-, 40-, 105-mm, etc.).

(3) Location accuracy.

(4) Equipment setup, calibration time and corresponding man-hour requirements.

(5) Survey time and corresponding man-hour requirements.

(6) Reacquisition/resurvey time and man-hour requirements (if any).

(7) Downtime due to system malfunctions and maintenance requirements.

1.3 STANDARD AND NONSTANDARD INERT ORDNANCE TARGETS

The standard and nonstandard ordnance items emplaced in the test areas are listed in Table 1. Standardized targets are members of a set of specific ordnance items that have identical properties to all other items in the set (caliber, configuration, size, weight, aspect ratio, material, filler, magnetic remanence, and nomenclature). Nonstandard targets are inert ordnance items having properties that differ from those in the set of standardized targets.

TABLE 1. INERT ORDNANCE TARGETS

Standard Type	Nonstandard (NS)
20-mm Projectile M55	20-mm Projectile M55
	20-mm Projectile M97
40-mm Grenades M385	40-mm Grenades M385
40-mm Projectile MKII Bodies	40-mm Projectile M813
BDU-28 Submunition	
BLU-26 Submunition	
M42 Submunition	
57-mm Projectile APC M86	
60-mm Mortar M49A3	60-mm Mortar (JPG)
	60-mm Mortar M49
2.75-inch Rocket M230	2.75-inch Rocket M230
	2.75-inch Rocket XM229
MK 118 ROCKEYE	
81-mm Mortar M374	81-mm Mortar (JPG)
	81-mm Mortar M374
105-mm Heat Rounds M456	
105-mm Projectile M60	105-mm Projectile M60
155-mm Projectile M483A1	155-mm Projectile M483A
	500-lb Bomb

JPG = Jefferson Proving Ground
HEAT = high-explosive, antitank

SECTION 2. DEMONSTRATION

2.1 DEMONSTRATOR INFORMATION

2.1.1 Demonstrator Point of Contact (POC) and Address

POC: Mr. Rob Siegel
617-964-7070 (extention: 262)

Address: GEO-CENTERS, Inc.
7 Wells Avenue
Newton, MA 02459

2.1.2 System Description (provided by demonstrator)

The Simultaneous EM and Magnetometry system (multisensor STOLS) (fig. 1) is a towed vehicular array developed by GEO-CENTERS and Corps of Engineers - Huntsville Center (CEHNC) with funding from ESTCP under project UX-0208. The system simultaneously collects both total field magnetometer data and EM61 data on a single towed platform. GEO-CENTERS' existing Surface Towed Ordnance Location System (STOLS) was used as a host system; STOLS' custom-fabricated aluminum dune buggy with a low-magnetic self-signature, magnetometers, differential GPS, sensors, computers, and tractor-trailer for transportation were reused. The new Simultaneous Electromagnetic (EM) and Magnetometry system augments STOLS with interleaved sampling electronics that allow EM61 coils to be physically located on the same platform as the magnetometers without corrupting the magnetometer data. The electronics monitor the rising edge of the 75 Hz transmit pulse from the EM61, wait 8 ms for the pulse to die down, sample the magnetometers for 5 ms, then wait for the next transmit pulse and repeat the cycle. Data acquired last month at McKinley Test Range (Redstone Arsenal, Huntsville) show that magnetometer data quality with the EM system switched on is commeasurate with magnetometer data quality when the EM system is switched off. Magnetometer, EM61, and GPS data are acquired in a single file.

Along with new interleaved sampling, electronics is a new proof-of-concept non-metallic tow platform to host both the EM61 coils and the magnetometers in a low-noise environment. Constructed almost entirely from fiberglass, the only metallic components on the platform are the axles, the hub, and a small number of aluminum pop rivets. The wheels are composite. Even the tires have had the metal beads removed. Total metallic mass has been reduced by over 99 percent by weight as compared to the original aluminum STOLS tow platform. Certain key structural locations have been reinforced with marine-grade plywood. The proof-of-concept platform was recently fielded successfully for a prove-out at McKinley Test Range. It should be noted that the platform was designed to fit into the existing budget for the ESTCP project, but was not designed for commercial surveys: it has no suspension, is speed-limited, and may not survive a fielding over rugged terrain without sustaining structural damage.

Five Geometrics 822A magnetometers updating and outputting at 75 Hz are deployed at 1/2 meter spacing. The magnetometers are 10 feet behind the tow vehicle. Three 1/2 meter Geonics EM61 coils (upper and lower) internally updating at 75 Hz and outputting at 10 Hz are deployed in a master/slave configuration on the rear of the platform, 8 feet behind the magnetometers, also at 1/2 meter spacing. The center line of the middle three magnetometers is coincident with the center line of the three EM61 coils. Both the magnetometers and the lower EM61 coils are mounted on pivots so they can swing up if they encounter an obstacle while moving forward.

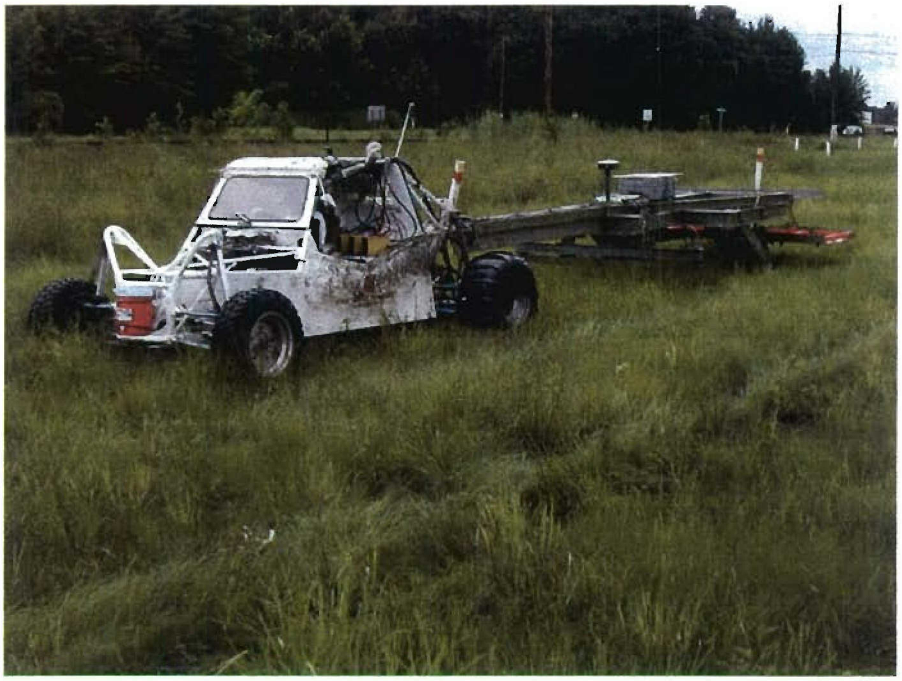

Figure 1. Demonstrator's system, STOLS/towed array.

2.1.3 <u>Data Processing Description (provided by demonstrator)</u>

Custom Unix-based data processing software is used to process the file containing the magnetometer, EM61, and GPS data. The GPS updates are automatically examined, and any jumps that could not occur at a nominal vehicle speed are flagged, allowing the operator to manually correct them. Sensor heading is calculated using smoothed position updates.

Magnetometer and EM61 data are then processed separately as they require different corrections. For the magnetometer data, the reference magnetometer recording the ambient variations of the Earth's magnetic field is time-correlated, then subtracted off. The data are then directionally divided into passes acquired in uniform directions (that is, north-going, south-going, west-going, and east-going, or whatever set of directions are used for the survey site). For each major direction, an independent set of sensor offsets are calculated and are then applied to that set of data to background-level the sensors and remove streaks in the image. A site-wide offset may also be applied if the reference magnetometer is over geology with a background different than that of the survey site.

EM61 background is not directionally dependent, but EM61 data are background-leveled individually by file to account for drift that may occur file-to-file.

Once the background-leveling corrections have been determined, data are processed. Adjacent 1-Hz GPS updates are used to position the sensor array at the beginning and at the end of each second. From there, each sensor on the array can be positioned at each of its updates. An array is set up by the data processing software at the 10 cm cell spacing, and each sensor update is positioned into the appropriate cell in the array. A nearest-neighbor-inverse-distance-squared interpolation is used to fill in the inter-sensor spacing regardless of the direction of travel. The interpolated image is then displayed on the screen for analysis.

Analysis of the magnetometer is performed using a nonlinear least-squares match to a model of a point dipole with adjustable angles. Outputs from the model are object location, depth, magnetic moment, angle of incidence, and angle of orientation. On the basis of magnetic moment, an estimate is made of object size. For objects that do not resemble point dipoles because they are either too weak or too spatially extended, the object's location can be pinpointed using the mouse. An optional comment field may be added to each target.

Simultaneous viewing and analysis of the simultaneously-collected magnetometer and EM data is obtained by running two linked copies of the data processing software. Once linked, panning, zooming, and scrolling in one set of data automatically pans, zooms, and scrolls in the other set. Drawing a region of interest in one set of data automatically draws the same region in the other set.

Data output is available in a variety of formats, including raw, corrected (navigation corrected and background-leveled), and interpolated.

2.1.4 <u>Data Submission Format</u>

Data were submitted for scoring in accordance with data submission protocols outlined in the Standardized UXO Technology Demonstration Site Handbook. These submitted data are not included in this report in order to protect ground truth information.

2.1.5 <u>Demonstrator Quality Assurance (QA) and Quality Control (QC) (provided by demonstrator)</u>

Overview of QC. The following QC steps are taken:

- Coordinates of the control monument over which to set up the base GPS station are obtained before deploying to the survey site. These coordinates are obtained in both latitude and longitude (WGS84) as well as the rectangular coordinate system used for final data submission (preferably UTM WGS84 meters) so we can verify that coordinates can be correctly converted between these two coordinate systems.

- The system is set up using checklists for the vehicle and platform, GPS, and diurnal variation stations.

- GPS data, magnetometer data, and EM61 data are all numerically displayed in a Windows program on the data acquisition computer. These numbers are all visually inspected prior to survey data acquisition, and at the beginning and end of each survey line.

- The six line test required by CEHNC is performed.

Overview of QA. The following QA steps are taken:

- Data are processed and imaged in the field immediately after survey operations to ensure that the data are of nominal quality.

- Any available control points, such as grid corner coordinates, are overlaid to ensure that the GPS was properly set up and that there are no coordinate offsets.

- Reference data are displayed to ensure that there are not unphysical spikes or dropouts.

- During processing, GPS data are viewed and corrected if necessary.

- Magnetometer data are reference-corrected.

- Magnetometer data are background-leveled using a correction specific to the direction of travel.

- EM61 data are background-leveled individually for each data file to mitigate the effects of drift.

- After data are converted to the desired data output format (e.g., American Standard Code for Information Interchange (ASCII), comma-delimited .dat files), these file are read back into the Unix-based data processing software, processed, and viewed.

2.1.6 <u>Additional Records</u>

The following record(s) by this vendor can be accessed via the Internet as MicroSoft Word documents at www.uxotestsites.org. The Blind Grid counterpart to this report is Scoring Record No. 290.

2.2 APG SITE INFORMATION

2.2.1 Location

The APG Standardized Test Site is located within a secured range area of the Aberdeen Area of APG. The Aberdeen Area of APG is located approximately 30 miles northeast of Baltimore at the northern end of the Chesapeake Bay. The Standardized Test Site encompasses 17 acres of upland and lowland flats, woods, and wetlands.

2.2.2 Soil Type

According to the soils survey conducted for the entire area of APG in 1998, the test site consists primarily of Elkton Series type soil (ref 2). The Elkton Series consists of very deep, slowly permeable, poorly drained soils. These soils formed in silty aeolin sediments and the underlying loamy alluvial and marine sediments. They are on upland and lowland flats and in depressions of the Mid-Atlantic Coastal Plain. Slopes range from 0 to 2 percent.

ERDC conducted a site-specific analysis in May of 2002 (ref 3). The results basically matched the soil survey mentioned above. Seventy percent of the samples taken were classified as silty loam. The majority (77 percent) of the soil samples had a measured water content between 15- and 30-percent with the water content decreasing slightly with depth.

For more details concerning the soil properties at the APG test site, go to www.uxotestsites.org on the web to view the entire soils description report.

2.2.3 Test Areas

A description of the test site areas at APG is included in Table 2.

TABLE 2. TEST SITE AREAS

Area	Description
Calibration Grid	Contains 14 standard ordnance items buried in six positions at various angles and depths to allow demonstrator to calibrate their equipment.
Blind Test Grid	Contains 400 grid cells in a 0.2-hectare (0.5 acre) site. The center of each grid cell contains ordnance, clutter or nothing.
Open Field	A 4-hectare (10-acre) site containing open areas, dips, ruts and obstructions that challenge platform systems or hand held detectors. The challenges include a gravel road, wet areas and trees. The vegetation height varies from 15 to 25 cm.

SECTION 3. FIELD DATA

3.1 DATE OF FIELD ACTIVITIES (4 through 6 August 2004)

3.2 AREAS TESTED/NUMBER OF HOURS

Areas tested and total number of hours operated at each site are summarized in Table 3.

TABLE 3. AREAS TESTED AND NUMBER OF HOURS

Area	Number of Hours
Calibration Lanes	0.75
Open Field	13.33

3.3 TEST CONDITIONS

3.3.1 Weather Conditions

An APG weather station located approximately one mile west of the test site was used to record average temperature and precipitation on a half hour basis for each day of operation. The temperatures listed in Table 4 represent the average temperature during field operations from 0700 to 1700 hours while precipitation data represents a daily total amount of rainfall. Hourly weather logs used to generate this summary are provided in Appendix B.

TABLE 4. TEMPERATURE/PRECIPITATION DATA SUMMARY

Date, 2004	Average Temperature, °F	Total Daily Precipitation, in.
4 August	84.55	0.06
5 August	72.91	0.03
6 August	66.7	0.00

3.3.2 Field Conditions

GEO-CENTERS surveyed the Open Field 4 and 5 August 2004. The Open Field had several muddy areas due to rain prior and during testing. Approximately 5-percent of the Open Field in the wet area could not be surveyed due to poor conditions. The vehicle was not able to traverse in these areas

3.3.3 Soil Moisture

Three soil probes were placed at various locations within the site to capture soil moisture data: Blind Grid, Calibration, Mogul, and Wooded areas. Measurements were collected in percent moisture and were taken twice daily (morning and afternoon) from five different soil depths (1 to 6 in., 6 to 12 in., 12 to 24 in., 24 to 36 in., and 36 to 48 in.) from each probe. Soil moisture logs are included in Appendix C.

3.4 FIELD ACTIVITIES

3.4.1 Setup/Mobilization

These activities included initial mobilization and daily equipment preparation and break down. A two-person crew took 6 hours and 15 minutes to perform the initial setup and mobilization. There was 55 minutes of daily equipment preparation and end of the day equipment break down lasted 35 minutes.

3.4.2 Calibration

No calibration activities occurred while surveying in the Open Field. GEO-CENTERS spent a total of 45 minutes in the Calibration Lanes, of which 15 minutes was spent collecting data.

3.4.3 Downtime Occasions

Occasions of downtime are grouped into five categories: equipment/data checks or equipment maintenance, equipment failure and repair, weather, Demonstration Site issues, or breaks/lunch. All downtime is included for the purposes of calculating labor costs (section 5) except for downtime due to Demonstration Site issues. Demonstration Site issues, while noted in the Daily Log, are considered nonchargeable downtime for the purposes of calculating labor costs and are not discussed. Breaks and lunches are discussed in this section and billed to the total Site Survey area.

3.4.3.1 Equipment/data checks, maintenance. Equipment/data checks and maintenance activities accounted for 55 minutes in the Open Field. GEO-CENTER had two data checks during the 55 minutes. GEO-CENTER also spent 1 hour and 10 minutes on breaks and lunches.

3.4.3.2 Equipment failure or repair. One equipment failure occurred in the Open Field. GEO-CENTER had a bad GPS satellite quality for 45 minutes on 4 August 2004. The situation rectified itself and no other problems occurred.

3.4.3.3 Weather. There were areas of standing water and mud in the Open Field. The weather on the survey days was generally warm and sunny.

3.4.4 Data Collection

GEO-CENTERS spent a total of 13 hours and 20 minutes in the Open Field, of which 9 hours was spent collecting data in the Open Field

3.4.5 Demobilization

The GEO-CENTERS survey crew went on to conduct a full demonstration of the site. Therefore, demobilization did not occur until 5 and 6 August 2004. On that day, it took the crew 3 hours and 45 minutes to break down and pack up their equipment.

3.5 PROCESSING TIME

GEO-CENTERS submitted the raw data from the demonstration activities on the last day of the demonstration, as required. The scoring submittal data was also provided within the required 30-day timeframe.

3.6 DEMONSTRATOR'S FIELD PERSONNEL

Rob Siegel, GEO-CENTERS, principle investigator
Roger J. Young, project lead from CEHNC, contracted by GEO-CENTERS
Alan Crandall, U.S. Environmental, contracted by GEO-CENTERS

3.7 DEMONSTRATOR'S FIELD SURVEYING METHOD

GEO-CENTER surveyed the Open Field in a linear fashion. The team started in the southwest corner and surveyed in a south/north direction. GEO-CENTER avoided the saturated areas and surveyed what they could at the end of the demonstration. It was estimated that 5 percent of the Open Field was too wet for surveying.

3.8 SUMMARY OF DAILY LOGS

Daily logs capture all field activities during this demonstration and are located in Appendix D. Activities pertinent to this specific demonstration are indicated in highlighted text.

SECTION 4. TECHNICAL PERFORMANCE RESULTS

4.1 ROC CURVES USING ALL ORDNANCE CATEGORIES

Figure 2, 4, and 6 shows the probability of detection for the response stage (P_d^{res}) and the discrimination stage (P_d^{disc}) versus their respective probability of false positive for the EM sensor(s), MAG sensor(s) and combined EM/MAG picks respectively. Figure 3, 5, and 7 shows both probabilities plotted against their respective background alarm rate. Both figures use horizontal lines to illustrate the performance of the demonstrator at two demonstrator-specified points: at the system noise level for the response stage, representing the point below which targets are not considered detectable, and at the demonstrator's recommended threshold level for the discrimination stage, defining the subset of targets the demonstrator would recommend digging based on discrimination. Note that all points have been rounded to protect the ground truth.

The overall ground truth is composed of ferrous and non-ferrous anomalies. Due to limitations of the magnetometer, the non-ferrous items cannot be detected. Therefore, the ROC curves presented in figures 4 and 5 of this section are based on the subset of the ground truth that is solely made up of ferrous anomalies.

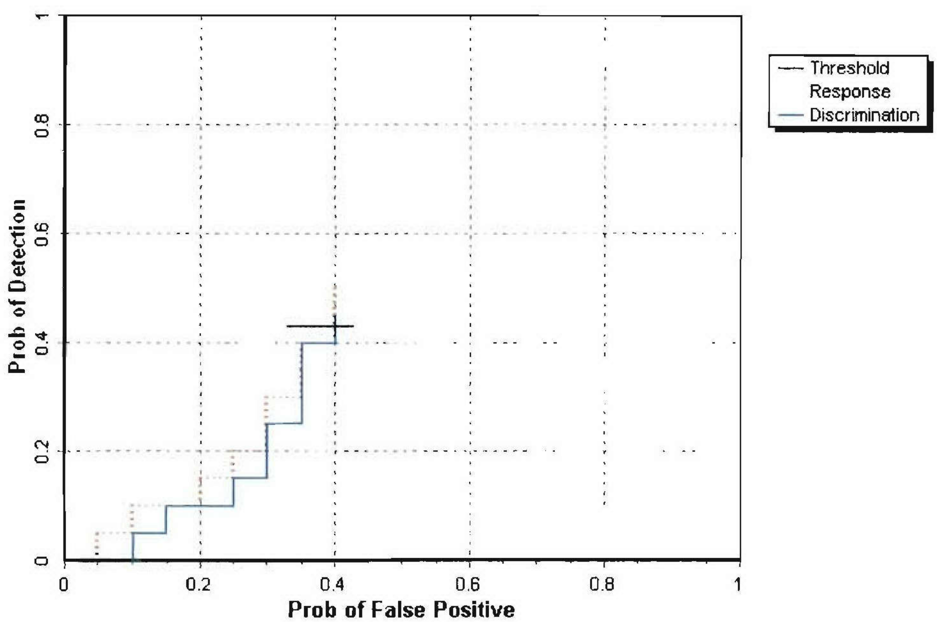

Figure 2. EM Sensor open field probability of detection for response and discrimination stages versus their respective probability of false positive over all ordnance categories combined.

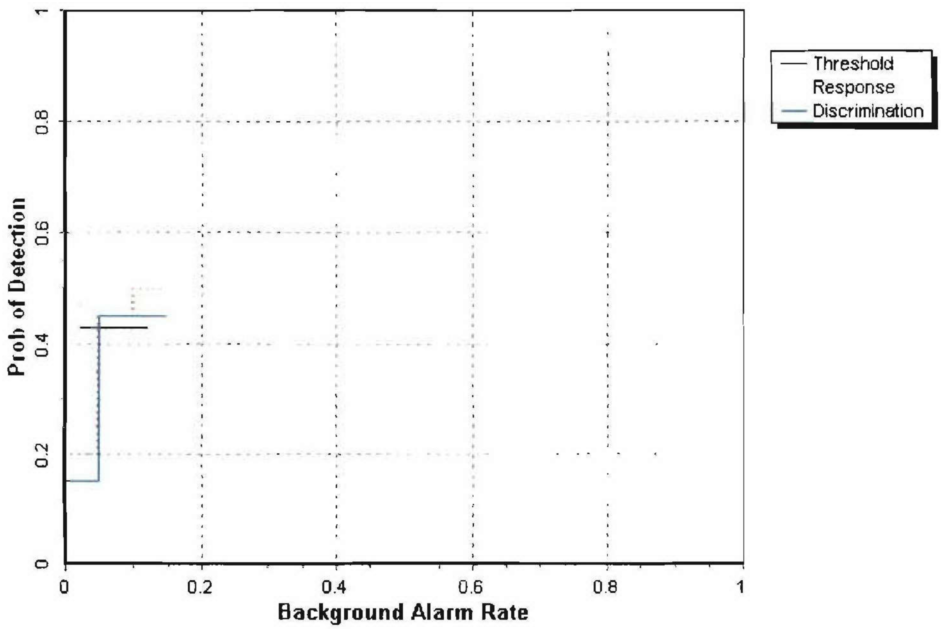

Figure 3. EM Sensor open field probability of detection for response and discrimination stages versus their respective background alarm rate over all ordnance categories combined.

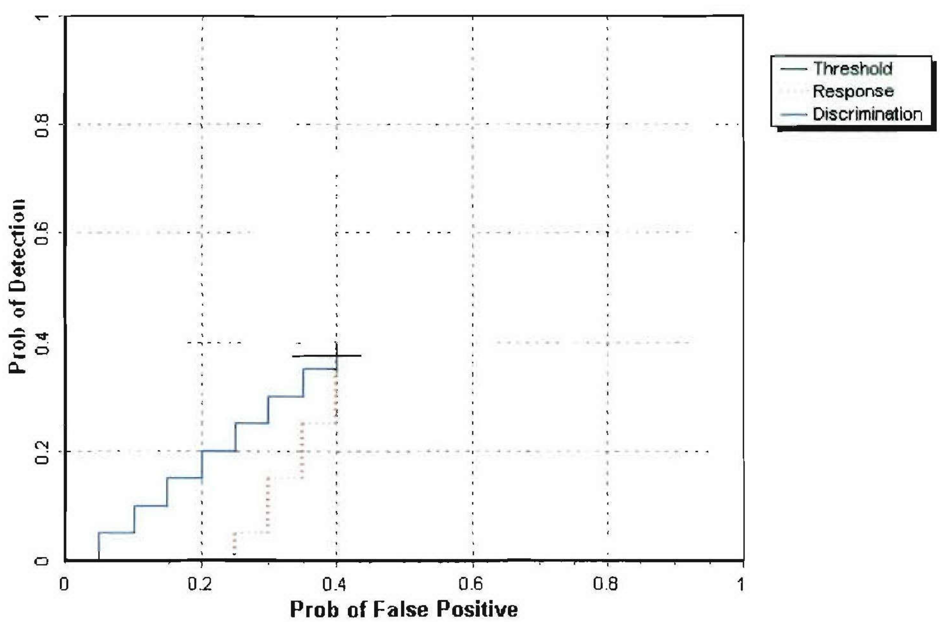

Figure 4. MAG Sensor open field probability of detection for response and discrimination stages versus their respective probability of false positive over all ordnance categories combined.

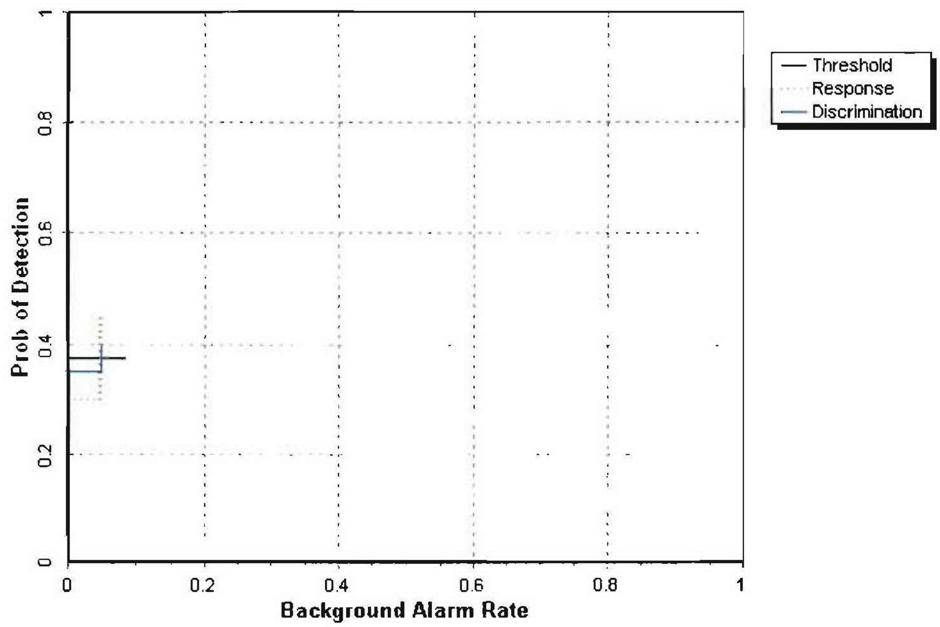

Figure 5. MAG Sensor open field probability of detection for response and discrimination stages versus their respective background alarm rate over all ordnance categories combined.

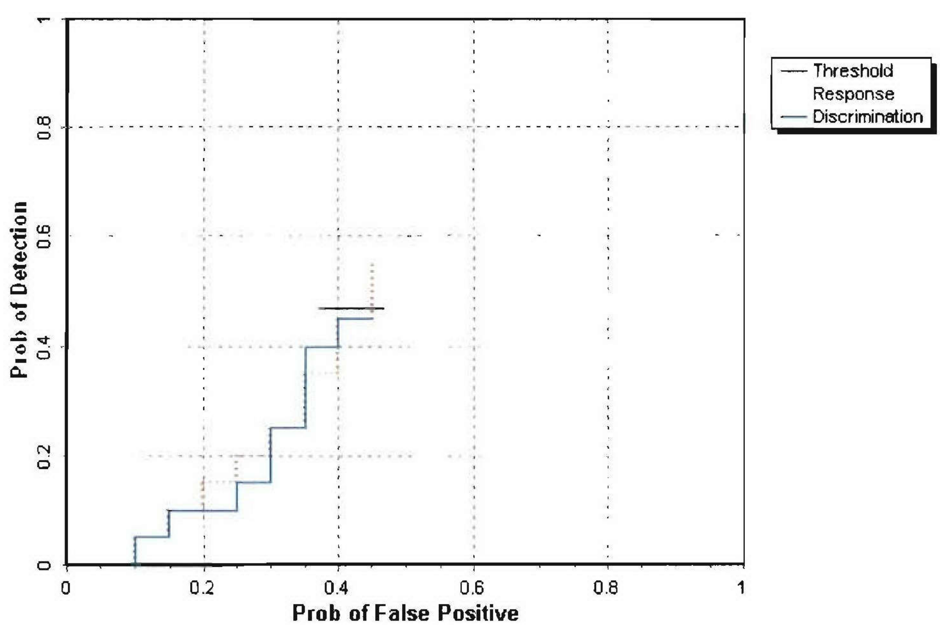

Figure 6. Combined Sensor open field probability of detection for response and discrimination stages versus their respective probability of false positive over all ordnance categories combined.

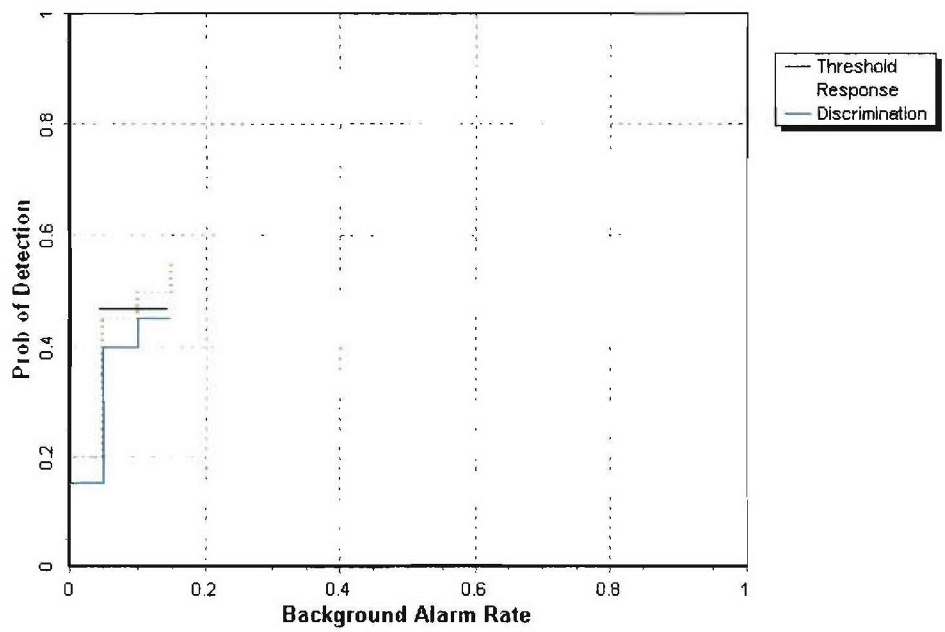

Figure 7. Combined Sensor open field probability of detection for response and discrimination stages versus their respective background alarm rate over all ordnance categories combined.

4.2 ROC CURVES USING ORDNANCE LARGER THAN 20 MM

Figure 8, 10, and 12 shows the probability of detection for the response stage (P_d^{res}) and the discrimination stage (P_d^{disc}) versus their respective probability of false positive when only targets larger than 20 mm are scored for the EM sensor(s), MAG sensor(s) and Combined EM/MAG picks respectively. Figure 9, 11, and 13 shows both probabilities plotted against their respective background alarm rate. Both figures use horizontal lines to illustrate the performance of the demonstrator at two demonstrator-specified points: at the system noise level for the response stage, representing the point below which targets are not considered detectable, and at the demonstrator's recommended threshold level for the discrimination stage, defining the subset of targets the demonstrator would recommend digging based on discrimination. Note that all points have been rounded to protect the ground truth.

The overall ground truth is composed of ferrous and non-ferrous anomalies. Due to limitations of the magnetometer, the non-ferrous items cannot be detected. Therefore, the ROC curves presented in figures 10 and 11 of this section are based on the subset of the ground truth that is solely made up of ferrous anomalies.

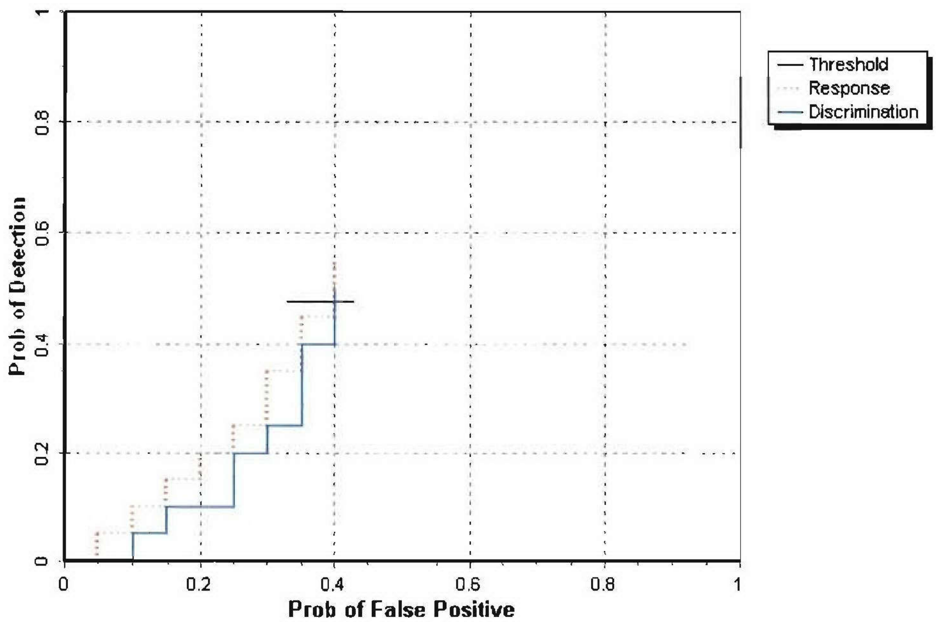

Figure 8. EM Sensor open field probability of detection for response and discrimination stages versus their respective probability of false positive for all ordnance larger than 20 mm.

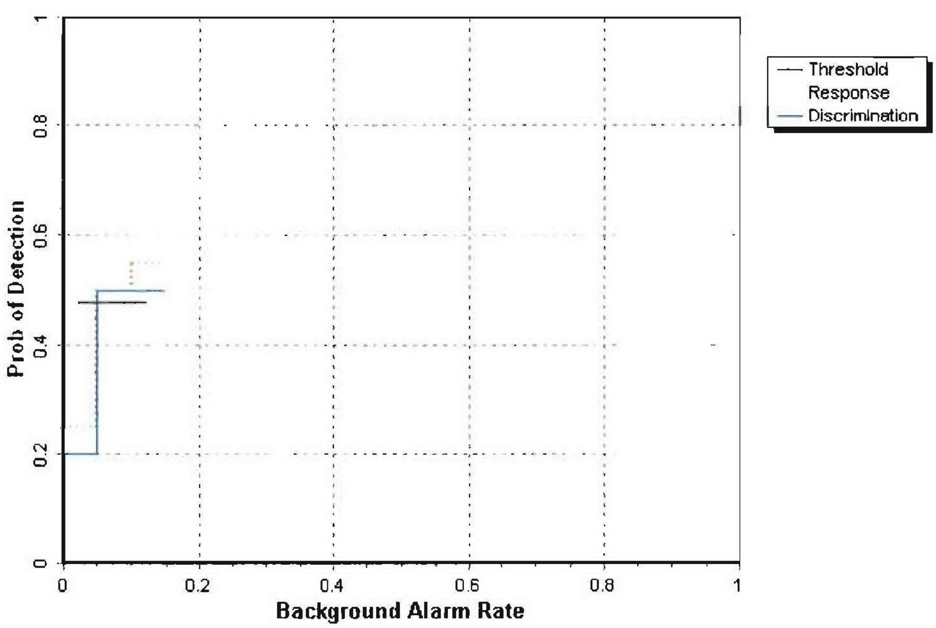

Figure 9. EM Sensor open field probability of detection for response and discrimination stages versus their respective background alarm rate for all ordnance larger than 20 mm.

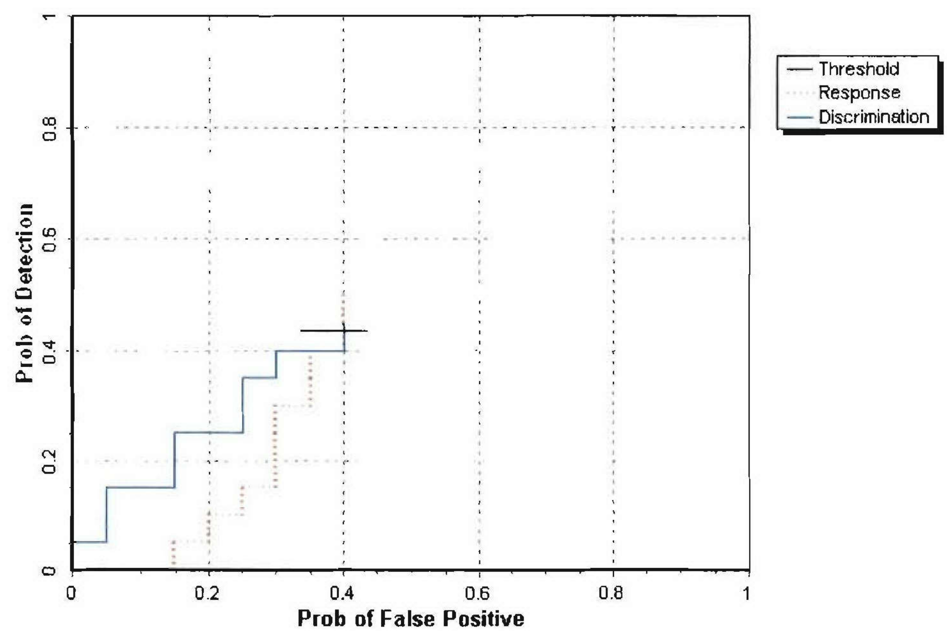

Figure 10. MAG Sensor open field probability of detection for response and discrimination stages versus their respective probability of false positive for all ordnance larger than 20 mm.

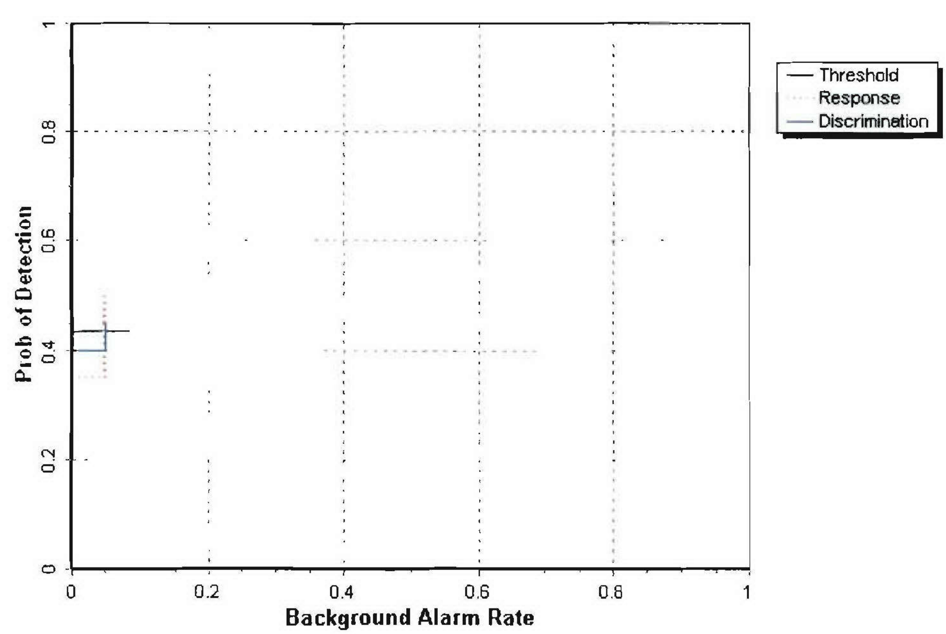

Figure 11. MAG Sensor open field probability of detection for response and discrimination stages versus their respective background alarm rate for all ordnance larger than 20 mm.

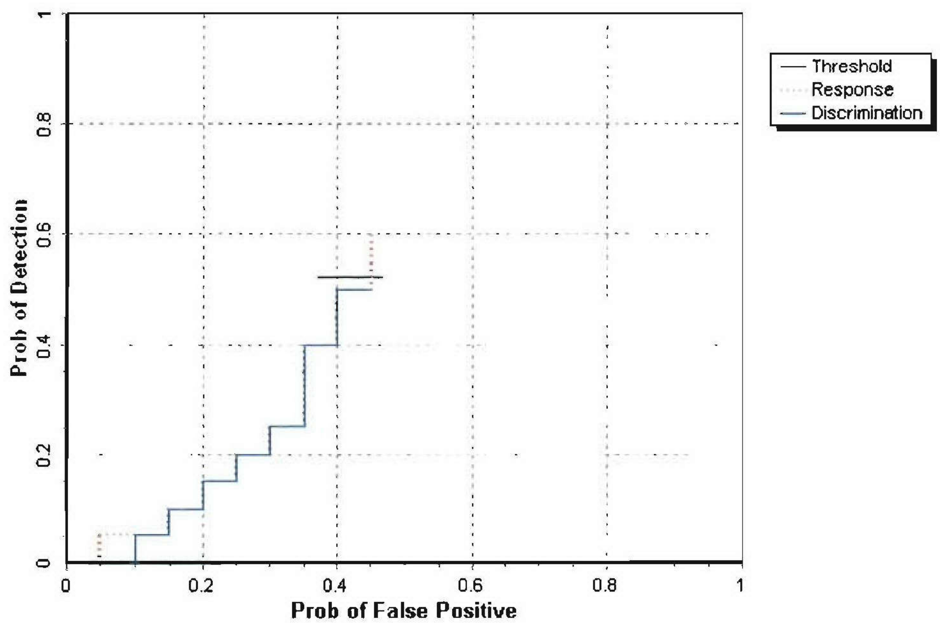

Figure 12. Combined Sensor open field probability of detection for response and discrimination stages versus their respective probability of false positive for all ordnance larger than 20 mm.

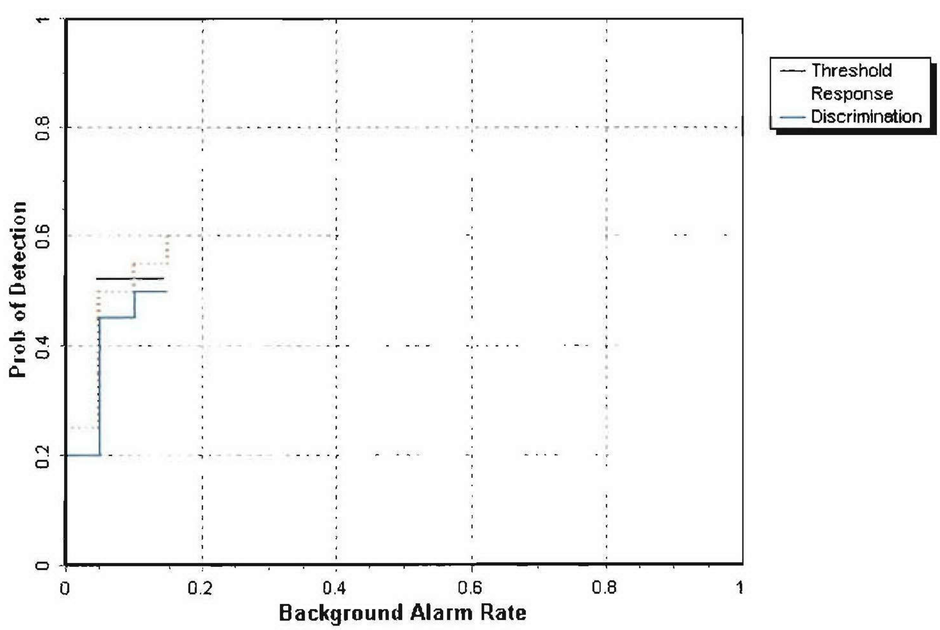

Figure 13. Combined Sensor open field probability of detection for response and discrimination stages versus their respective background alarm rate for all ordnance larger than 20 mm.

4.3 PERFORMANCE SUMMARIES

Results for the Open Field test broken out by sensor type, size, depth and nonstandard ordnance are presented in Tables 5a, b, and c (for cost results, see section 5). Results by size and depth include both standard and nonstandard ordnance. The results by size show how well the demonstrator did at detecting/discriminating ordnance of a certain caliber range (see app A for size definitions). The results are relative to the number of ordnance items emplaced. Depth is measured from the geometric center of anomalies.

The RESPONSE STAGE results are derived from the list of anomalies above the demonstrator-provided noise level. The results for the DISCRIMINATION STAGE are derived from the demonstrator's recommended threshold for optimizing UXO field cleanup by minimizing false digs and maximizing ordnance recovery. The lower 90-percent confidence limit on probability of detection and P_{fp} was calculated assuming that the number of detections and false positives are binomially distributed random variables. All results in Table 5 have been rounded to protect the ground truth. However, lower confidence limits were calculated using actual results.

The overall ground truth is composed of ferrous and non-ferrous anomalies. Due to limitations of the magnetometer, the non-ferrous items cannot be detected. Therefore, the summary presented in Table 5b is split exhibiting results based on the subset of the ground truth that is solely the ferrous anomalies and the full ground truth for comparison purposes.

All other tables presented in this section are based on scoring against the ferrous only ground truth. The response stage noise level and recommended discrimination stage threshold values are provided by the demonstrator.

TABLE 5A. SUMMARY OF OPEN FIELD RESULTS FOR THE STOLS/TOWED ARRAY (EM SENSOR)

Metric	Overall	Standard	Nonstandard	By Size			By Depth, m		
				Small	Medium	Large	< 0.3	0.3 to <1	>= 1
RESPONSE STAGE									
P_d	0.50	0.55	0.40	0.40	0.55	0.65	0.55	0.45	0.40
P_d Low 90% Conf	0.46	0.51	0.35	0.35	0.47	0.55	0.52	0.40	0.30
P_d Upper 90% Conf	0.53	0.61	0.46	0.46	0.60	0.71	0.62	0.53	0.47
P_{fp}	0.40	-	-	-	-	-	0.35	0.45	0.55
P_{fp} Low 90% Conf	0.39	-	-	-	-	-	0.32	0.44	0.38
P_{fp} Upper 90% Conf	0.43	-	-	-	-	-	0.38	0.50	0.74
BAR	0.15	-	-	-	-	-	-	-	-
DISCRIMINATION STAGE									
P_d	0.45	0.50	0.35	0.30	0.50	0.60	0.45	0.45	0.30
P_d Low 90% Conf	0.39	0.43	0.30	0.23	0.44	0.52	0.41	0.38	0.24
P_d Upper 90% Conf	0.46	0.53	0.41	0.33	0.57	0.68	0.52	0.50	0.40
P_{fp}	0.40	-	-	-	-	-	0.30	0.45	0.50
P_{fp} Low 90% Conf	0.36	-	-	-	-	-	0.28	0.42	0.32
P_{fp} Upper 90% Conf	0.40	-	-	-	-	-	0.33	0.48	0.68
BAR	0.05	-	-	-	-	-	-	-	-

Response Stage Noise Level: -0.22
Recommended Discrimination Stage Threshold: 3.00

TABLE 5b. SUMMARY OF OPEN FIELD RESULTS FOR THE STOLS/TOWED ARRAY (MAG SENSOR)

Ferrous Only Ground Truth									
				By Size			By Depth, m		
Metric	Overall	Standard	Nonstandard	Small	Medium	Large	< 0.3	0.3 to <1	>= 1
RESPONSE STAGE									
P_d	0.45	0.45	0.40	0.20	0.50	0.70	0.45	0.45	0.45
P_d Low 90% Conf	0.40	0.42	0.34	0.17	0.42	0.62	0.37	0.39	0.36
P_d Upper 90% Conf	0.48	0.52	0.46	0.28	0.54	0.77	0.49	0.52	0.53
P_{fp}	0.40	-	-	-	-	-	0.30	0.50	0.70
P_{fp} Low 90% Conf	0.38	-	-	-	-	-	0.28	0.45	0.50
P_{fp} Upper 90% Conf	0.43	-	-	-	-	-	0.34	0.51	0.84
BAR	0.05	-	-	-	-	-	-	-	-
DISCRIMINATION STAGE									
P_d	0.40	0.40	0.35	0.10	0.45	0.65	0.35	0.40	0.40
P_d Low 90% Conf	0.34	0.34	0.30	0.08	0.37	0.56	0.29	0.34	0.31
P_d Upper 90% Conf	0.41	0.44	0.42	0.18	0.49	0.72	0.40	0.47	0.48
P_{fp}	0.40	-	-	-	-	-	0.30	0.45	0.65
P_{fp} Low 90% Conf	0.36	-	-	-	-	-	0.27	0.43	0.43
P_{fp} Upper 90% Conf	0.41	-	-	-	-	-	0.33	0.49	0.79
BAR	0.05	-	-	-	-	-	-	-	-

Full Ground Truth									
				By Size			By Depth, m		
Metric	Overall	Standard	Nonstandard	Small	Medium	Large	< 0.3	0.3 to <1	>= 1
RESPONSE STAGE									
P_d	0.45	0.50	0.40	0.30	0.50	0.70	0.45	0.50	0.45
P_d Low 90% Conf	0.42	0.43	0.36	0.24	0.43	0.63	0.38	0.42	0.36
P_d Upper 90% Conf	0.49	0.53	0.47	0.34	0.56	0.78	0.49	0.54	0.53
P_{fp}	0.35	-	-	-	-	-	0.35	0.30	0.20
P_{fp} Low 90% Conf	0.32	-	-	-	-	-	0.34	0.29	0.07
P_{fp} Upper 90% Conf	0.36	-	-	-	-	-	0.39	0.34	0.37
BAR	0.05	-	-	-	-	-	-	-	-
DISCRIMINATION STAGE									
P_d	0.35	0.35	0.30	0.50	0.25	0.05	0.50	0.25	0.10
P_d Low 90% Conf	0.30	0.30	0.26	0.46	0.22	0.03	0.43	0.20	0.06
P_d Upper 90% Conf	0.37	0.39	0.37	0.57	0.33	0.12	0.54	0.31	0.18
P_{fp}	0.35	-	-	-	-	-	0.45	0.35	0.05
P_{fp} Low 90% Conf	0.35	-	-	-	-	-	0.40	0.30	0.01
P_{fp} Upper 90% Conf	0.39	-	-	-	-	-	0.46	0.36	0.22
BAR	0.05	-	-	-	-	-	-	-	-

Response Stage Noise Level: 2.81
Recommended Discrimination Stage Threshold: 1.00

TABLE 5c. SUMMARY OF OPEN FIELD RESULTS FOR THE STOLS/TOWED ARRAY (COMBINED EM/MAG RESULTS)

Metric	Overall	Standard	Nonstandard	By Size			By Depth, m		
				Small	Medium	Large	< 0.3	0.3 to <1	>= 1
RESPONSE STAGE									
P_d	0.55	0.60	0.45	0.40	0.55	0.70	0.60	0.50	0.45
P_d Low 90% Conf	0.50	0.53	0.41	0.37	0.51	0.62	0.54	0.43	0.37
P_d Upper 90% Conf	0.57	0.62	0.52	0.48	0.63	0.77	0.64	0.56	0.55
P_{fp}	0.45	-	-	-	-	-	0.40	0.50	0.75
P_{fp} Low 90% Conf	0.43	-	-	-	-	-	0.35	0.49	0.56
P_{fp} Upper 90% Conf	0.48	-	-	-	-	-	0.41	0.55	0.89
BAR	0.15	-	-	-	-	-	-	-	-
DISCRIMINATION STAGE									
P_d	0.45	0.50	0.40	0.30	0.55	0.65	0.50	0.50	0.35
P_d Low 90% Conf	0.43	0.47	0.33	0.27	0.48	0.56	0.44	0.43	0.28
P_d Upper 90% Conf	0.51	0.57	0.45	0.38	0.61	0.72	0.55	0.55	0.45
P_{fp}	0.40	-	-	-	-	-	0.30	0.50	0.75
P_{fp} Low 90% Conf	0.40	-	-	-	-	-	0.29	0.47	0.56
P_{fp} Upper 90% Conf	0.44	-	-	-	-	-	0.35	0.53	0.89
BAR	0.10	-	-	-	-	-	-	-	-

Response Stage Noise Level: -6.50
Recommended Discrimination Stage Threshold: 2.99

Note: The recommended discrimination stage threshold values are provided by the demonstrator.

4.4 EFFICIENCY, REJECTION RATES, AND TYPE CLASSIFICATION (All results based on combined EM/MAG data set)

Efficiency and rejection rates are calculated to quantify the discrimination ability at specific points of interest on the ROC curve: (1) at the point where no decrease in P_d is suffered (i.e., the efficiency is by definition equal to one) and (2) at the operator selected threshold. These values are reported in Table 6.

TABLE 6. EFFICIENCY AND REJECTION RATES

	Efficiency (E)	False Positive Rejection Rate	Background Alarm Rejection Rate
At Operating Point	0.88	0.08	0.47
With No Loss of P_d	1.00	0.02	0.02

At the demonstrator's recommended setting, the ordnance items that were detected and correctly discriminated were further scored on whether their correct type could be identified (table 7). Correct type examples include "20-mm projectile, 105-mm HEAT Projectile, and 2.75-inch Rocket". A list of the standard type declaration required for each ordnance item was provided to demonstrators prior to testing. For example, the standard type for the three example items are 20mmP, 105H, and 2.75in, respectively.

TABLE 7. CORRECT TYPE CLASSIFICATION OF TARGETS CORRECTLY DISCRIMINATED AS UXO

Size	Percentage Correct
Small	NA
Medium	NA
Large	NA
Overall	NA

4.5 LOCATION ACCURACY

The mean location error and standard deviations appear in Table 8. These calculations are based on average missed depth for ordnance correctly identified in the discrimination stage. Depths are measured from the closest point of the ordnance to the surface. For the Blind Grid, only depth errors are calculated, since (X, Y) positions are known to be the centers of each grid square.

TABLE 8. MEAN LOCATION ERROR AND STANDARD DEVIATION (M)

	Mean	Standard Deviation
Northing	0.00	0.21
Easting	-0.01	0.19
Depth	0.03	0.23

SECTION 5. ON-SITE LABOR COSTS

A standardized estimate for labor costs associated with this effort was calculated as follows: the first person at the test site was designated "supervisor", the second person was designated "data analyst", and the third and following personnel were considered "field support". Standardized hourly labor rates were charged by title: supervisor at $95.00/hour, data analyst at $57.00/hour, and field support at $28.50/hour.

Government representatives monitored on-site activity. All on-site activities were grouped into one of ten categories: initial setup/mobilization, daily setup/stop, calibration, collecting data, downtime due to break/lunch, downtime due to equipment failure, downtime due to equipment/data checks or maintenance, downtime due to weather, downtime due to demonstration site issue, or demobilization. See Appendix D for the daily activity log. See section 3.4 for a summary of field activities.

The standardized cost estimate associated with the labor needed to perform the field activities is presented in Table 9. Note that calibration time includes time spent in the Calibration Lanes as well as field calibrations. "Site survey time" includes daily setup/stop time, collecting data, breaks/lunch, downtime due to equipment/data checks or maintenance, downtime due to failure, and downtime due to weather.

TABLE 9. ON-SITE LABOR COSTS

	No. People	Hourly Wage	Hours	Cost
INITIAL SETUP				
Supervisor	1	$95.00	6.25	$593.75
Data Analyst	1	57.00	6.25	356.25
Field Support	0	28.50	6.25	0.00
Subtotal				**$950.00**
CALIBRATION				
Supervisor	1	$95.00	0.75	$71.25
Data Analyst	1	57.00	0.75	42.75
Field Support	0	28.50	0.75	0.00
Subtotal				**$114.00**
SITE SURVEY				
Supervisor	1	$95.00	13.33	$1,266.35
Data Analyst	1	57.00	13.33	759.81
Field Support	0	28.50	13.33	0.00
Subtotal				**$2,026.16**

See notes at end of table.

TABLE 9 (CONT'D)

	No. People	Hourly Wage	Hours	Cost
DEMOBILIZATION				
Supervisor	1	$95.00	3.75	$356.25
Data Analyst	1	57.00	3.75	213.75
Field Support	0	28.50	3.75	0.00
Subtotal				**$570.00**
Total				**$3,660.16**

Notes: Calibration time includes time spent in the Calibration Lanes as well as calibration before each data run.

Site Survey time includes daily setup/stop time, collecting data, breaks/lunch, downtime due to system maintenance, failure, and weather.

SECTION 6. COMPARISON OF RESULTS TO BLIND GRID DEMONSTRATION (BASED ON COMBINED EM/MAG DATA SETS)

6.1 SUMMARY OF RESULTS FROM BLIND GRID DEMONSTRATION

Table 10 shows the results from the Blind Grid survey conducted prior to surveying the Open Field during the same site visit in August of 2004. Due to the system utilizing magnetometer type sensors, all results presented in the following section have been based on performance scoring against the ferrous only ground truth anomalies. For more details on the Blind Grid survey results reference section 2.1.6.

TABLE 10. SUMMARY OF BLIND GRID RESULTS FOR THE STOLS/TOWED ARRAY

Metric	Overall	Standard	Nonstandard	By Size			By Depth, m		
				Small	Medium	Large	< 0.3	0.3 to <1	>= 1
RESPONSE STAGE									
P_d	0.70	0.80	0.65	0.75	0.70	0.80	0.85	0.80	0.20
P_d Low 90% Conf	0.65	0.69	0.50	0.63	0.55	0.55	0.75	0.67	0.08
P_d Upper 90% Conf	0.79	0.86	0.74	0.83	0.79	0.95	0.92	0.89	0.42
P_{fp}	0.80	-	-	-	-	-	0.80	0.75	1.00
P_{fp} Low 90% Conf	0.73	-	-	-	-	-	0.71	0.66	0.63
P_{fp} Upper 90% Conf	0.85	-	-	-	-	-	0.88	0.85	1.00
P_{ba}	0.10	-	-	-	-	-	-	-	-
DISCRIMINATION STAGE									
P_d	0.40	0.45	0.30	0.20	0.60	0.70	0.30	0.65	0.20
P_d Low 90% Conf	0.33	0.35	0.20	0.11	0.45	0.45	0.18	0.52	0.08
P_d Upper 90% Conf	0.47	0.55	0.44	0.29	0.70	0.88	0.39	0.77	0.42
P_{fp}	0.65	-	-	-	-	-	0.60	0.60	1.00
P_{fp} Low 90% Conf	0.56	-	-	-	-	-	0.51	0.48	0.63
P_{fp} Upper 90% Conf	0.69	-	-	-	-	-	0.71	0.70	1.00
P_{ba}	0.00	-	-	-	-	-	-	-	-

6.2 COMPARISON OF ROC CURVES USING ALL ORDNANCE CATEGORIES

Figure 6 shows P_d^{res} versus the respective P_{fp} over all ordnance categories. Figure 7 shows P_d^{disc} versus their respective P_{fp} over all ordnance categories. Figure 7 uses horizontal lines to illustrate the performance of the demonstrator at the recommended discrimination threshold levels, defining the subset of targets the demonstrator would recommend digging based on discrimination. The ROC curves in this section are a sole reflection of the ferrous only survey.

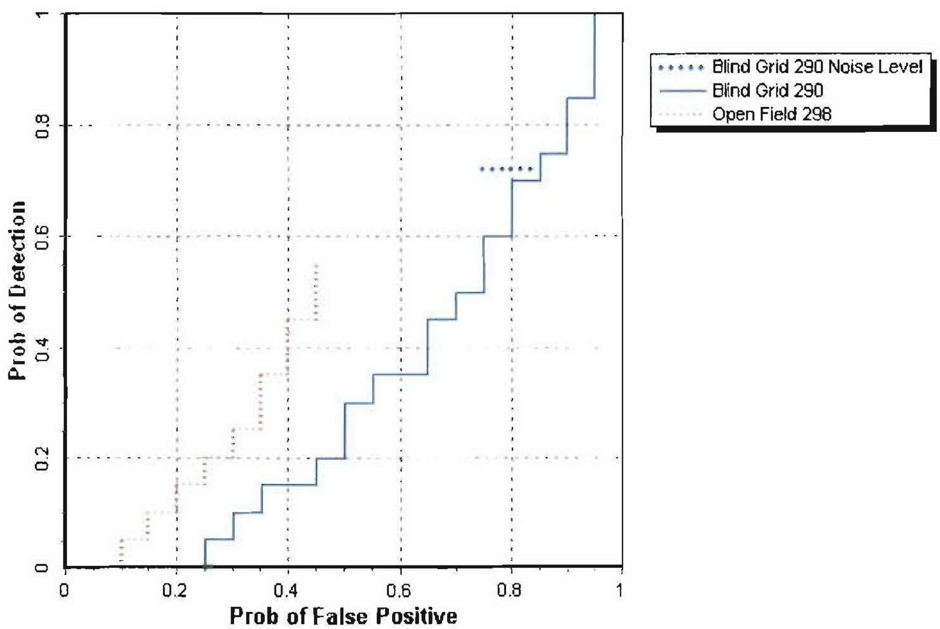

Figure 6. STOLS/towed array dual mode P_d^{res} stages versus the respective P_{fp} over all ordnance categories combined.

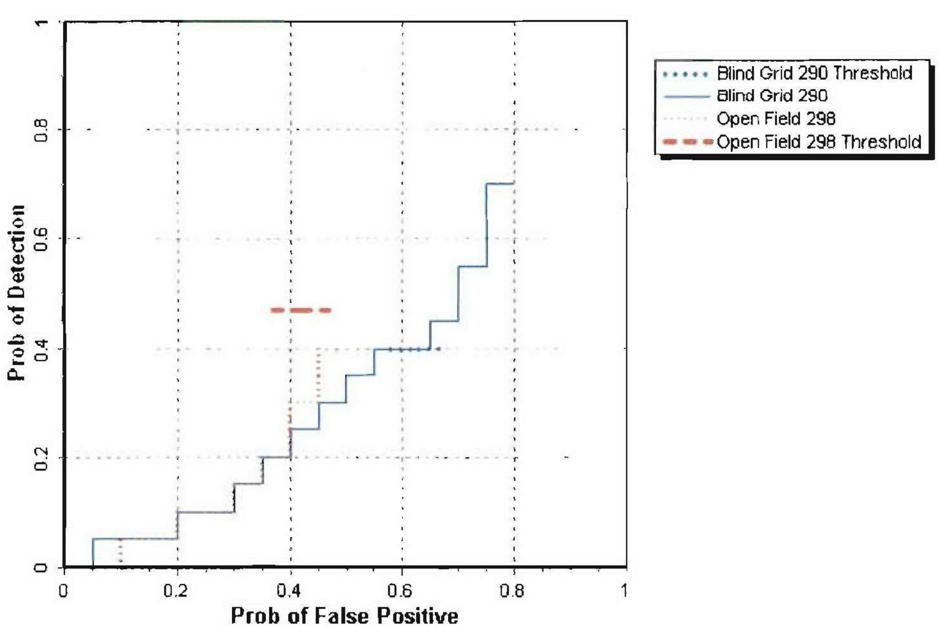

Figure 7. STOLS/towed array dual mode P_d^{disc} versus the respective P_{fp} over all ordnance categories combined.

6.3 COMPARISON OF ROC CURVES USING ORDNANCE LARGER THAN 20 MM

Figure 8 shows the P_d^{res} versus the respective probability of P_{fp} over ordnance larger than 20 mm. Figure 9 shows P_d^{disc} versus the respective P_{fp} over ordnance larger than 20 mm. Figure 9 uses horizontal lines to illustrate the performance of the demonstrator at the recommended discrimination threshold levels, defining the subset of targets the demonstrator would recommend digging based on discrimination.

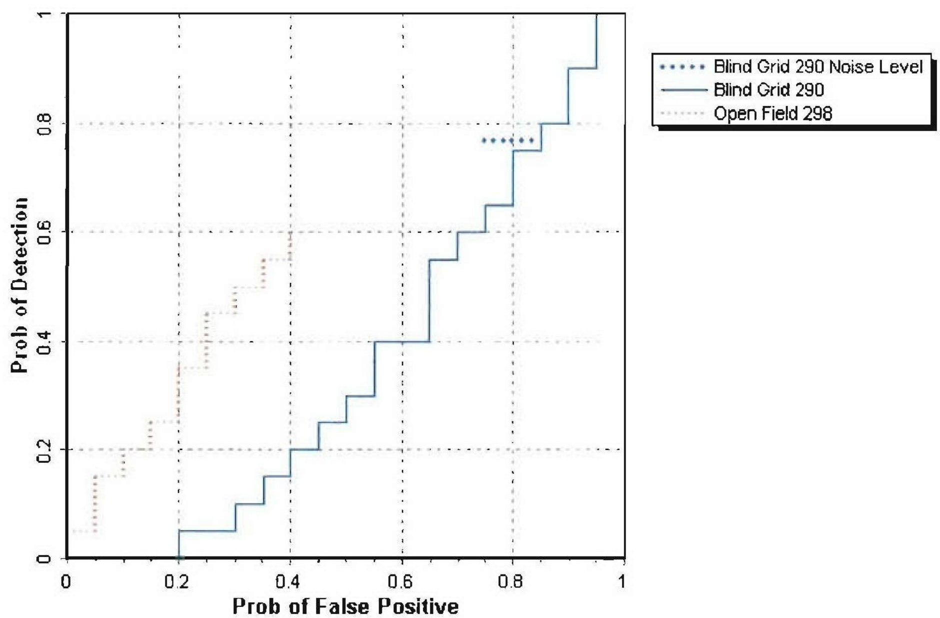

Figure 8. STOLS/towed array dual mode P_d^{res} versus the respective P_{fp} for ordnance larger than 20 mm.

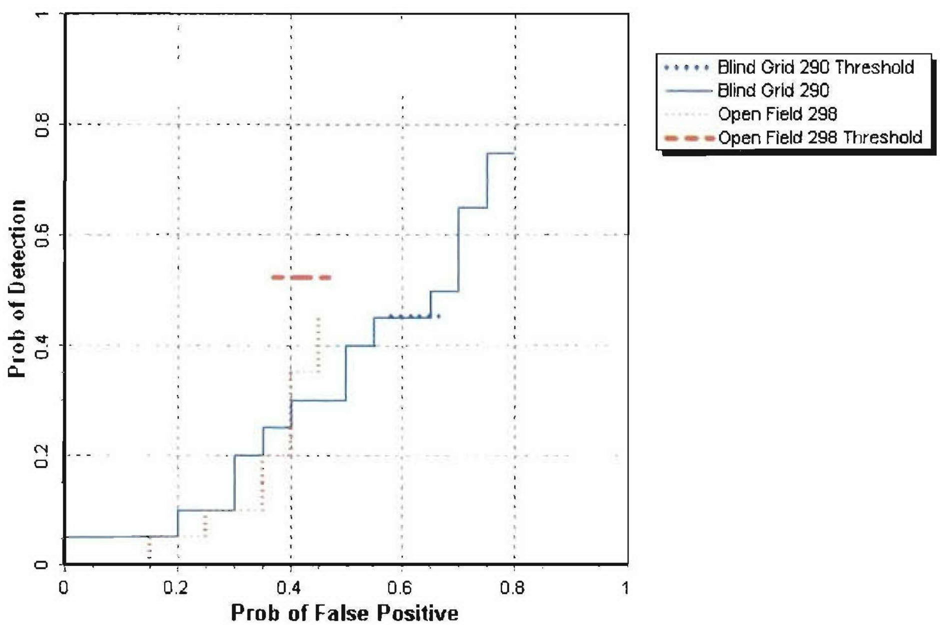

Figure 9. STOLS/towed array dual mode P_d^{disc} versus the respective P_{fp} for ordnance larger than 20 mm.

6.4 STATISTICAL COMPARISONS

Statistical Chi-square significance tests were used to compare results between the Blind Grid and Open Field scenarios. The intent of the comparison is to determine if the feature introduced in each scenario has a degrading effect on the performance of the sensor system. However, any modifications in the UXO sensor system during the test, like changes in the processing or changes in the selection of the operating threshold, will also contribute to performance differences.

The Chi-square test for comparison between ratios was used at a significance level of 0.05 to compare Blind Grid to Open Field with regard to P_d^{res}, P_d^{disc}, P_{fp}^{res} and P_{fp}^{disc}, Efficiency and Rejection Rate. These results are presented in Table 11. A detailed explanation and example of the Chi-square application is located in Appendix A.

TABLE 11. CHI-SQUARE RESULTS - BLIND GRID VERSUS OPEN FIELD

Metric	Small	Medium	Large	Overall
P_d^{res}	Significant	Not Significant	Not Significant	Significant
P_d^{disc}	Not Significant	Not Significant	Not Significant	Not Significant
P_{fp}^{res}	Not Significant	Not Significant	Not Significant	Significant
P_{fp}^{disc}	-	-	-	Significant
Efficiency	-			Significant
Rejection rate	-	-	-	Significant

SECTION 7. APPENDIXES

APPENDIX A. TERMS AND DEFINITIONS

GENERAL DEFINITIONS

Anomaly: Location of a system response deemed to warrant further investigation by the demonstrator for consideration as an emplaced ordnance item.

Detection: An anomaly location that is within R_{halo} of an emplaced ordnance item.

Emplaced Ordnance: An ordnance item buried by the government at a specified location in the test site.

Emplaced Clutter: A clutter item (i.e., non-ordnance item) buried by the government at a specified location in the test site.

R_{halo}: A pre-determined radius about the periphery of an emplaced item (clutter or ordnance) within which a location identified by the demonstrator as being of interest is considered to be a response from that item. If multiple declarations lie within R_{halo} of any item (clutter or ordnance), the declaration with the highest signal output within the R_{halo} will be utilized. For the purpose of this program, a circular halo 0.5 meters in radius will be placed around the center of the object for all clutter and ordnance items less than 0.6 meters in length. When ordnance items are longer than 0.6 meters, the halo becomes an ellipse where the minor axis remains 1 meter and the major axis is equal to the length of the ordnance plus 1 meter.

Small Ordnance: Caliber of ordnance less than or equal to 40 mm (includes 20-mm projectile, 40-mm projectile, submunitions BLU-26, BLU-63, and M42).

Medium Ordnance: Caliber of ordnance greater than 40 mm and less than or equal to 81 mm (includes 57-mm projectile, 60-mm mortar, 2.75 in. Rocket, MK118 Rockeye, 81-mm mortar).

Large Ordnance: Caliber of ordnance greater than 81 mm (includes 105-mm HEAT, 105-mm projectile, 155-mm projectile, 500-pound bomb).

Shallow: Items buried less than 0.3 meter below ground surface.

Medium: Items buried greater than or equal to 0.3 meter and less than 1 meter below ground surface.

Deep: Items buried greater than or equal to 1 meter below ground surface.

Response Stage Noise Level: The level that represents the point below which anomalies are not considered detectable. Demonstrators are required to provide the recommended noise level for the Blind Grid test area.

Discrimination Stage Threshold: The demonstrator selected threshold level that they believe provides optimum performance of the system by retaining all detectable ordnance and rejecting the maximum amount of clutter. This level defines the subset of anomalies the demonstrator would recommend digging based on discrimination.

Binomially Distributed Random Variable: A random variable of the type which has only two possible outcomes, say success and failure, is repeated for n independent trials with the probability p of success and the probability 1-p of failure being the same for each trial. The number of successes x observed in the n trials is an estimate of p and is considered to be a binomially distributed random variable.

RESPONSE AND DISCRIMINATION STAGE DATA

The scoring of the demonstrator's performance is conducted in two stages. These two stages are termed the RESPONSE STAGE and DISCRIMINATION STAGE. For both stages, the probability of detection (P_d) and the false alarms are reported as receiver operating characteristic (ROC) curves. False alarms are divided into those anomalies that correspond to emplaced clutter items, measuring the probability of false positive (P_{fp}) and those that do not correspond to any known item, termed background alarms.

The RESPONSE STAGE scoring evaluates the ability of the system to detect emplaced targets without regard to ability to discriminate ordnance from other anomalies. For the RESPONSE STAGE, the demonstrator provides the scoring committee with the location and signal strength of all anomalies that the demonstrator has deemed sufficient to warrant further investigation and/or processing as potential emplaced ordnance items. This list is generated with minimal processing (e.g., this list will include all signals above the system noise threshold). As such, it represents the most inclusive list of anomalies.

The DISCRIMINATION STAGE evaluates the demonstrator's ability to correctly identify ordnance as such, and to reject clutter. For the same locations as in the RESPONSE STAGE anomaly list, the DISCRIMINATION STAGE list contains the output of the algorithms applied in the discrimination-stage processing. This list is prioritized based on the demonstrator's determination that an anomaly location is likely to contain ordnance. Thus, higher output values are indicative of higher confidence that an ordnance item is present at the specified location. For electronic signal processing, priority ranking is based on algorithm output. For other systems, priority ranking is based on human judgment. The demonstrator also selects the threshold that the demonstrator believes will provide "optimum" system performance, (i.e., that retains all the detected ordnance and rejects the maximum amount of clutter).

Note: The two lists provided by the demonstrator contain identical numbers of potential target locations. They differ only in the priority ranking of the declarations.

RESPONSE STAGE DEFINITIONS

Response Stage Probability of Detection (P_d^{res}): P_d^{res} = (No. of response-stage detections)/(No. of emplaced ordnance in the test site).

Response Stage False Positive (fp^{res}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Response Stage Probability of False Positive (P_{fp}^{res}): P_{fp}^{res} = (No. of response-stage false positives)/(No. of emplaced clutter items).

Response Stage Background Alarm (ba^{res}): An anomaly in a blind grid cell that contains neither emplaced ordnance nor an emplaced clutter item. An anomaly location in the open field or scenarios that is outside R_{halo} of any emplaced ordnance or emplaced clutter item.

Response Stage Probability of Background Alarm (P_{ba}^{res}): Blind Grid only: P_{ba}^{res} = (No. of response-stage background alarms)/(No. of empty grid locations).

Response Stage Background Alarm Rate (BAR^{res}): Open Field only: BAR^{res} = (No. of response-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{res}, P_{fp}^{res}, P_{ba}^{res}, and BAR^{res} are functions of t^{res}, the threshold applied to the response-stage signal strength. These quantities can therefore be written as $P_d^{res}(t^{res})$, $P_{fp}^{res}(t^{res})$, $P_{ba}^{res}(t^{res})$, and $BAR^{res}(t^{res})$.

DISCRIMINATION STAGE DEFINITIONS

Discrimination: The application of a signal processing algorithm or human judgment to response-stage data that discriminates ordnance from clutter. Discrimination should identify anomalies that the demonstrator has high confidence correspond to ordnance, as well as those that the demonstrator has high confidence correspond to nonordnance or background returns. The former should be ranked with highest priority and the latter with lowest.

Discrimination Stage Probability of Detection (P_d^{disc}): P_d^{disc} = (No. of discrimination-stage detections)/(No. of emplaced ordnance in the test site).

Discrimination Stage False Positive (fp^{disc}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Discrimination Stage Probability of False Positive (P_{fp}^{disc}): P_{fp}^{disc} = (No. of discrimination stage false positives)/(No. of emplaced clutter items).

Discrimination Stage Background Alarm (ba^{disc}): An anomaly in a blind grid cell that contains neither emplaced ordnance nor an emplaced clutter item. An anomaly location in the open field or scenarios that is outside R_{halo} of any emplaced ordnance or emplaced clutter item.

Discrimination Stage Probability of Background Alarm (P_{ba}^{disc}): P_{ba}^{disc} = (No. of discrimination-stage background alarms)/(No. of empty grid locations).

Discrimination Stage Background Alarm Rate (BAR^{disc}): BAR^{disc} = (No. of discrimination-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{disc}, P_{fp}^{disc}, P_{ba}^{disc}, and BAR^{disc} are functions of t^{disc}, the threshold applied to the discrimination-stage signal strength. These quantities can therefore be written as $P_d^{disc}(t^{disc})$, $P_{fp}^{disc}(t^{disc})$, $P_{ba}^{disc}(t^{disc})$, and $BAR^{disc}(t^{disc})$.

RECEIVER-OPERATING CHARACERISTIC (ROC) CURVES

ROC curves at both the response and discrimination stages can be constructed based on the above definitions. The ROC curves plot the relationship between P_d versus P_{fp} and P_d versus BAR or P_{ba} as the threshold applied to the signal strength is varied from its minimum (t_{min}) to its maximum (t_{max}) value.[1] Figure A-1 shows how P_d versus P_{fp} and P_d versus BAR are combined into ROC curves. Note that the "res" and "disc" superscripts have been suppressed from all the variables for clarity.

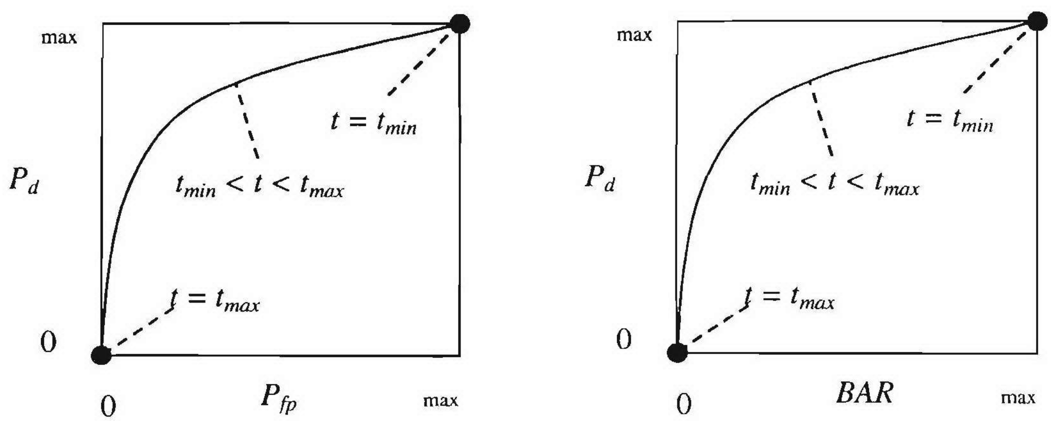

Figure A-1. ROC curves for open field testing. Each curve applies to both the response and discrimination stages.

[1]Strictly speaking, ROC curves plot the P_d versus P_{ba} over a pre-determined and fixed number of detection opportunities (some of the opportunities are located over ordnance and others are located over clutter or blank spots). In an open field scenario, each system suppresses its signal strength reports until some bare-minimum signal response is received by the system. Consequently, the open field ROC curves do not have information from low signal-output locations, and, furthermore, different contractors report their signals over a different set of locations on the ground. These ROC curves are thus not true to the strict definition of ROC curves as defined in textbooks on detection theory. Note, however, that the ROC curves obtained in the Blind Grid test sites are true ROC curves.

METRICS TO CHARACTERIZE THE DISCRIMINATION STAGE

The demonstrator is also scored on efficiency and rejection ratio, which measure the effectiveness of the discrimination stage processing. The goal of discrimination is to retain the greatest number of ordnance detections from the anomaly list, while rejecting the maximum number of anomalies arising from nonordnance items. The efficiency measures the amount of detected ordnance retained by the discrimination, while the rejection ratio measures the fraction of false alarms rejected. Both measures are defined relative to the entire response list, i.e., the maximum ordnance detectable by the sensor and its accompanying false positive rate or background alarm rate.

Efficiency (E): $E = P_d^{disc}(t^{disc})/P_d^{res}(t_{min}^{res})$; Measures (at a threshold of interest), the degree to which the maximum theoretical detection performance of the sensor system (as determined by the response stage tmin) is preserved after application of discrimination techniques. Efficiency is a number between 0 and 1. An efficiency of 1 implies that all of the ordnance initially detected in the response stage was retained at the specified threshold in the discrimination stage, t^{disc}.

False Positive Rejection Rate (R_{fp}): $R_{fp} = 1 - [P_{fp}^{disc}(t^{disc})/P_{fp}^{res}(t_{min}^{res})]$; Measures (at a threshold of interest), the degree to which the sensor system's false positive performance is improved over the maximum false positive performance (as determined by the response stage tmin). The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all emplaced clutter initially detected in the response stage were correctly rejected at the specified threshold in the discrimination stage.

Background Alarm Rejection Rate (R_{ba}):

Blind Grid: $R_{ba} = 1 - [P_{ba}^{disc}(t^{disc})/P_{ba}^{res}(t_{min}^{res})]$.
Open Field: $R_{ba} = 1 - [BAR^{disc}(t^{disc})/BAR^{res}(t_{min}^{res})]$).

Measures the degree to which the discrimination stage correctly rejects background alarms initially detected in the response stage. The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all background alarms initially detected in the response stage were rejected at the specified threshold in the discrimination stage.

CHI-SQUARE COMPARISON EXPLANATION:

The Chi-square test for differences in probabilities (or 2 x 2 contingency table) is used to analyze two samples drawn from two different populations to see if both populations have the same or different proportions of elements in a certain category. More specifically, two random samples are drawn, one from each population, to test the null hypothesis that the probability of event A (some specified event) is the same for both populations (ref 3).

A 2 x 2 contingency table is used in the Standardized UXO Technology Demonstration Site Program to determine if there is reason to believe that the proportion of ordnance correctly detected/discriminated by demonstrator X's system is significantly degraded by the more challenging terrain feature introduced. The test statistic of the 2 x 2 contingency table is the

Chi-square distribution with one degree of freedom. Since an association between the more challenging terrain feature and relatively degraded performance is sought, a one-sided test is performed. A significance level of 0.05 is chosen which sets a critical decision limit of 2.71 from the Chi-square distribution with one degree of freedom. It is a critical decision limit because if the test statistic calculated from the data exceeds this value, the two proportions tested will be considered significantly different. If the test statistic calculated from the data is less than this value, the two proportions tested will be considered not significantly different.

An exception must be applied when either a 0 or 100 percent success rate occurs in the sample data. The Chi-square test cannot be used in these instances. Instead, Fischer's test is used and the critical decision limit for one-sided tests is the chosen significance level, which in this case is 0.05. With Fischer's test, if the test statistic is less than the critical value, the proportions are considered to be significantly different.

Standardized UXO Technology Demonstration Site examples, where blind grid results are compared to those from the open field and open field results are compared to those from one of the scenarios, follow. It should be noted that a significant result does not prove a cause and effect relationship exists between the two populations of interest; however, it does serve as a tool to indicate that one data set has experienced a degradation in system performance at a large enough level than can be accounted for merely by chance or random variation. Note also that a result that is not significant indicates that there is not enough evidence to declare that anything more than chance or random variation within the same population is at work between the two data sets being compared.

Demonstrator X achieves the following overall results after surveying each of the three progressively more difficult areas using the same system (results indicate the number of ordnance detected divided by the number of ordnance emplaced):

	Blind Grid	Open Field	Moguls
P_d^{res}	100/100 = 1.0	8/10 = .80	20/33 = .61
P_d^{disc}	80/100 = 0.80	6/10 = .60	8/33 = .24

P_d^{res}: BLIND GRID versus OPEN FIELD. Using the example data above to compare probabilities of detection in the response stage, all 100 ordnance out of 100 emplaced ordnance items were detected in the blind grid while 8 ordnance out of 10 emplaced were detected in the open field. Fischer's test must be used since a 100 percent success rate occurs in the data. Fischer's test uses the four input values to calculate a test statistic of 0.0075 that is compared against the critical value of 0.05. Since the test statistic is less than the critical value, the smaller response stage detection rate (0.80) is considered to be significantly less at the 0.05 level of significance. While a significant result does not prove a cause and effect relationship exists between the change in survey area and degradation in performance, it does indicate that the detection ability of demonstrator X's system seems to have been degraded in the open field relative to results from the blind grid using the same system.

P_d^{disc}: BLIND GRID versus OPEN FIELD. Using the example data above to compare probabilities of detection in the discrimination stage, 80 out of 100 emplaced ordnance items were correctly discriminated as ordnance in blind grid testing while 6 ordnance out of 10 emplaced were correctly discriminated as such in open field-testing. Those four values are used to calculate a test statistic of 1.12. Since the test statistic is less than the critical value of 2.71, the two discrimination stage detection rates are considered to be not significantly different at the 0.05 level of significance.

P_d^{res}: OPEN FIELD versus MOGULS. Using the example data above to compare probabilities of detection in the response stage, 8 out of 10 and 20 out of 33 are used to calculate a test statistic of 0.56. Since the test statistic is less than the critical value of 2.71, the two response stage detection rates are considered to be not significantly different at the 0.05 level of significance.

P_d^{disc}: OPEN FIELD versus MOGULS. Using the example data above to compare probabilities of detection in the discrimination stage, 6 out of 10 and 8 out of 33 are used to calculate a test statistic of 2.98. Since the test statistic is greater than the critical value of 2.71, the smaller discrimination stage detection rate is considered to be significantly less at the 0.05 level of significance. While a significant result does not prove a cause and effect relationship exists between the change in survey area and degradation in performance, it does indicate that the ability of demonstrator X to correctly discriminate seems to have been degraded by the mogul terrain relative to results from the flat open field using the same system.

APPENDIX B. DAILY WEATHER LOGS

TABLE B-1. WEATHER LOG

Date & Time	Average Temperature, °F	Total Precipitation, in.
08/03/2004		
0700	74.1	0
0800	77.2	0
0900	79.6	0
1000	81.8	0
1100	83.6	0
1200	84.5	0
1300	84.7	0
1400	86.7	0
1500	86.8	0
1600	87.5	0
1700	86.3	0
08/04/2004		
0700	76.2	0
0800	78.6	0
0900	81.2	0
1000	83.5	0
1100	84.9	0
1200	85.9	0
1300	87.7	0
1400	88.6	0
1500	87.9	0
1600	87.8	0
1700	87.8	0

Date & Time	Average Temperature, °F	Total Precipitation, in.
08/05/2004		
0700	71.1	0
0800	69.9	0
0900	70.4	0
1000	72.1	0
1100	72.9	0
1200	72.2	0
1300	72.9	0
1400	73.9	0
1500	74.7	0
1600	75.8	0
1700	76.1	0
08/06/2005		
0700	61.6	0
0800	64.1	0
0900	66.1	0
1000	67.9	0
1100	69.8	0
1200	70.7	0

Date: 8/5/2004
Times: 0800 hours, 1600 hours

Probe Location	Layer, in.	AM Reading, %	PM Reading, %
Wet Area	0 to 6	65.4	No Readings Taken
	6 to 12	75.8	
	12 to 24	79.1	
	24 to 36	55.5	
	36 to 48	52.8	
Wooded Area	0 to 6	No Readings Taken	No Readings Taken
	6 to 12		
	12 to 24		
	24 to 36		
	36 to 48		
Open Area	0 to 6	22.0	No Readings Taken
	6 to 12	6.9	
	12 to 24	19.0	
	24 to 36	26.1	
	36 to 48	52.8	
Calibration Lanes	0 to 6	No Readings Taken	No Readings Taken
	6 to 12		
	12 to 24		
	24 to 36		
	36 to 48		
Blind Grid/Moguls	0 to 6	No Readings Taken	No Readings Taken
	6 to 12		
	12 to 24		
	24 to 36		
	36 to 48		

APPENDIX D. DAILY ACTIVITY LOGS

Date	No. of People	Area Tested	Status Start Time	Status Stop Time	Duration, min	Operational Status	OP Stat Code	Operational Status - Comments	Track Method	Track Method=Other Explain	Pattern	Field Conditions	
8/3/2004	2	CALIBRATION LANE	1010	1635	375	INITIAL MOBILIZATION	2	INITIAL MOBILIZATION	GPS	NA	LINEAR	SUNNY	MUDDY
8/3/2004	2	CALIBRATION LANE	1635	1650	15	COLLECT DATA	4	COLLECT DATA	GPS	NA	LINEAR	SUNNY	MUDDY
8/3/2004	2	CALIBRATION LANE	1650	1720	30	DAILY START STOP	3	EQUIPMENT BREAKDOWN	GPS	NA	LINEAR	SUNNY	MUDDY
8/4/2004	2	BLIND TEST GRID	740	940	120	DAILY START STOP	3	SET UP OPERATIONS	GPS	NA	LINEAR	SUNNY	MUDDY
8/420/04	2	BLIND TEST GRID	940	1000	20	COLLECT DATA	4	COLLECT DATA	GPS	NA	LINEAR	SUNNY	MUDDY
8/4/2004	2	BLIND TEST GRID	1000	1010	10	DOWNTIME MAINTENANCE CHECK	7	DATA CHECK	GPS	NA	LINEAR	SUNNY	MUDDY
8/4/2004	2	OPEN FIELD	1010	1155	105	COLLECT DATA	4	COLLECT DATA	GPS	NA	LINEAR	SUNNY	MUDDY
8/4/2004	2	OPEN FIELD	1155	1305	70	LUNCH/BREAK	5	LUNCH/BREAK	GPS	NA	LINEAR	SUNNY	MUDDY
8/4/2004	2	OPEN FIELD	1305	1425	80	COLLECT DATA	4	COLLECT DATA	GPS	NA	LINEAR	SUNNY	MUDDY
8/4/2004	2	OPEN FIELD	1425	1510	45	EQUIPMENT FAILURE	6	BAD SATELLITE QUALITY	GPS	NA	LINEAR	SUNNY	MUDDY
8/4/2004	2	OPEN FIELD	1510	1705	115	COLLECT DATA	4	COLLECT DATA	GPS	NA	LINEAR	SUNNY	MUDDY
8/4/2004	2	OPEN FIELD	1705	1740	35	DAILY START STOP	3	EQUIPMENT BREAKDOWN	GPS	NA	LINEAR	SUNNY	MUDDY

Note: Activities pertinent to this specific demonstration are indicated in highlighted text.

Date	No. of People	Area Tested	Status Start Time	Status Stop Time	Duration, min	Operational Status	OP Stat Code	Operational Status - Comments	Track Method	Track Method=Other Explain	Pattern	Field Conditions	
8/5/2004	2	OPEN FIELD	745	840	55	DAILY START STOP	3	SET UP OPERATIONS	GPS	NA	LINEAR	CLOUDY	MUDDY
8/5/2004	2	OPEN FIELD	840	1005	85	COLLECT DATA	4	COLLECT DATA	GPS	NA	LINEAR	CLOUDY	MUDDY
8/5/2004	2	OPEN FIELD	1005	1040	35	DOWNTIME MAINTENANCE CHECK	7	DATA CHECK	GPS	NA	LINEAR	CLOUDY	MUDDY
8/5/2004	2	OPEN FIELD	1040	1200	80	COLLECT DATA	4	COLLECT DATA	GPS	NA	LINEAR	CLOUDY	MUDDY
8/5/2004	2	OPEN FIELD	1200	1220	20	DOWNTIME MAINTENANCE CHECK	7	DATA CHECK	GPS	NA	LINEAR	CLOUDY	MUDDY
8/5/2004	2	OPEN FIELD	1220	1335	75	COLLECT DATA	4	COLLECT DATA	GPS	NA	LINEAR	CLOUDY	MUDDY
8/5/2004	2	OPEN FIELD	1545	1715	90	DEMOBILIZATION	10	DEMOBILIZATION	GPS	NA	LINEAR	CLOUDY	MUDDY
8/6/2004	2	OPEN FIELD	745	1000	135	DEMOBILIZATION	10	DEMOBILIZATION	GPS	NA	LINEAR	CLOUDY	MUDDY
8/5/2004	2	ACTIVE SITE	1220	1335	75	COLLECT DATA	4	COLLECT DATA	GPS	NA	LINEAR	CLOUDY	MUDDY

Note: Activities pertinent to this specific demonstration are indicated in highlighted text.

APPENDIX E. REFERENCES

1. Standardized UXO Technology Demonstration Site Handbook, DTC Project No. 8-CO-160-000-473, Report No. ATC-8349, March 2002.

2. Aberdeen Proving Ground Soil Survey Report, October 1998.

3. Data Summary, UXO Standardized Test Site: APG Soils Description, May 2002.

4. Yuma Proving Ground Soil Survey Report, May 2003.

APPENDIX F. ABBREVIATIONS

AEC	=	U.S. Army Environmental Center
APG	=	Aberdeen Proving Ground
ASCII	=	American Standard Code for Information Interchange.
ATC	=	U.S. Army Aberdeen Test Center
CEHNC	=	Corps of Engineers - Huntsville Center
EM	=	electromagnetic
EMI	=	electromagnetic interference
EMIS	=	Electromagnetic Induction Spectroscopy
ERDC	=	U.S. Army Corps of Engineers Engineering Research and Development Center
ESTCP	=	Environmental Security Technology Certification Program
EQT	=	Army Environmental Quality Technology Program
GPS	=	Global Positioning System
HEAT	=	high-explosive, antitank
JPG	=	Jefferson Proving Ground
POC	=	point of contact
QA	=	quality assurance
QC	=	quality control
ROC	=	receiver-operating characteristic
SERDP	=	Strategic Environmental Research and Development Program
STOLS	=	Surface Towed Ordnance Location System
UXO	=	unexploded ordnance
YPG	=	U.S. Army Yuma Proving Ground

APPENDIX G. DISTRIBUTION LIST

DTC Project No. 8-CO-160-UXO-021

Addressee	No. of Copies
Commander U.S. Army Environmental Center ATTN: SFIM-AEC-ATT (Mr. George Robitaille) Aberdeen Proving Ground, MD 21010-5401	2
GEO-CENTERS, Inc. ATTN: (Mr. Rob Siegel) 7 Wells Avenue Newton, MA 02459	1
SERDP/ESTCP ATTN: (Ms. Anne Andrews) 901 North Stuart Street, Suite 303 Arlington, VA 22203	1
Commander U.S. Army Aberdeen Test Center ATTN: CSTE-DTC-SL-E (Mr. Larry Overbay)	1
(Library)	1
CSTE-DTC-AT-CS-R Aberdeen Proving Ground, MD 21005-5059	1
Defense Technical Information Center 8725 John J. Kingman Road, STE 0944 Fort Belvoir, VA 22060-6218	2

Secondary distribution is controlled by Commander, U.S. Army Environmental Center, ATTN: SFIM-AEC-ATT.